1960

Multiple factor analysis.

Phillips, Jack B.

Monterey, California: U.S. Naval Postgraduate School

MULTIPLE FACTOR ANALYSIS

JACK B. PHILLIPS

UNITED STATES
NAVAL POSTGRADUATE SCHOOL

THESIS

MULTIPLE FACTOR ANALYSIS

BY

JACK B. PHILLIPS

MAJOR, UNITED STATES MARINE CORPS

1960

12ND P2238 (1-59)

MULTIPLE FACTOR ANALYSIS

* * * * *

Jack B. Phillips

MULTIPLE FACTOR ANALYSIS

by

Jack B. Phillips
//
Major, United States Marine Corps

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE

United States Naval Postgraduate School
Monterey, California

1 9 6 0

[illegible]
1980
[illegible]

~~Thesis~~
[illegible]

MULTIPLE FACTOR ANALYSIS

by

Jack B. Phillips

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

from the

United States Naval Postgraduate School

ABSTRACT

Multiple factor analysis has been used extensively to explore mental measurements and allied subjects in psychology. Recently, it has been used with some success in exploring other problem areas of science. The concept of multiple factor analysis is discussed, a basic mathematical model is formulated and two methods of factoring are described in detail, both from the theoretical as well as the computational standpoint. The problems of rotations, rank of the correlation matrix and communalities are discussed.

The writer wishes to express his appreciation for the assistance given him by Professors Charles C. Torrance and Jack C. Borsting of the U. S. Naval Postgraduate School and Mr. Ray Twery of the Standford Research Institute.

PREFACE

This paper was undertaken with the objective of providing a sound basis for the study of multiple factor analysis by a person whose mathematical background includes elementary matrix theory and a working knowledge of statistics. Therefore, it is written for the novice factorist and not the expert.

Two types of footnotes are employed throughout the paper. A footnote enclosed in parantheses (1) refers to the general bibliography, while one without parantheses refers to a footnote in the normal sense.

The writer wishes to thank Miss Beth Millet of the U. S. Naval Postgraduate School for her clerical assistance and careful proofreading of this paper.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

Factor analysis is distinguished from most other branches of statistics by having been developed mainly by men who were not statisticians. Generally speaking, the main developments are due to psychologists. For this reason, factor analysis has developed a language and technique of its own.(8)

Although factor analysis was developed for the primary purpose of identifying the principal dimensions or categories of mentality, the methods are general so that they have been found useful for investigations in other sciences.(4)

When a particular domain is to be investigated by means of individual differences, one can proceed in one of two ways: invent a hypothesis regarding the processes that underlie the differences, and then set up an experiment to test that hypothesis; or, if no promising hypothesis is available, the domain may be represented in terms of a set of measurements or numerical indices, and then perform a factor analysis. The analysis may reveal an underlying order which would be of great assistance in formulating the scientific concepts concerning that particular domain.

The characteristic of factor analysis that gives it some value as a general scientific method is the second approach described above. That method may reveal the underlying order of a domain without first postulating it in the form of a hypothesis. In this sense, factor analysis has its principal usefulness at the borderlines of science, especially in those domains where basic concepts are lacking and where effective experimentation has been difficult to design.(8)

Basic concept of factor analysis.

Individuals have combinations of causal agents that produce combinations of effects. The objectivity of this concept is not defended, but the concept conforms with simple intuition and provides a framework that facilitates comprehension and analysis of logical relations.

An effect is that which is measured by some type of measuring instrument under some particular conditions. A measuring instrument may be physical (e.g., ammeter); it may be a question on a questionnaire, or it may be a human mind rendering a judgment.

In psychology, a set of measurements of a set of effects is called a test. But because the word "test" has other connotations and because a distinction must be made between a particular effect and a measured value of that effect, the terms "effect" and "measured effect" will be used, or, in mathematical language, "variable" and "observed value of a variable," or generally "observation."

> The experimental situation and the nature of the data on which the technique of factor analysis can be successfully employed may be stated as follows: Each of the measurements on an individual has a linear regression on a common set of a few hypothetical variables or factors. (6)

Factor analysis is a methodology for discovering a possible set of "causes" (factors) required to account for the observed values of n effects (variables) in p individuals. It makes no attempt to discern the metaphysical nature of the "causes."

CHAPTER 2

BASIC STATISTICS

Let X_{ij}, $(i=1,\cdots,p;\ j=1,\cdots,n)$ represent a set of n observations on each of p variables, let X be the matrix of such observations with typical element X_{ij}, and let X_i be a row vector of observations on the ith variable. Then, the mean of X_i is

$$\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij}. \tag{2-1}$$

Reducing the X_{ij} by $\bar{X}_i$, we define a deviate to be

$$x_{ij} = X_{ij} - \bar{X}_i . \tag{2-2}$$

The sample standard deviation of the vector, X_i, is defined by

$$S(X_i) = \sqrt{\frac{\sum_{j=1}^{n} x_{ij}^2}{n}} . \tag{2-3}$$

By taking the sample standard deviation for the vector, X_i, as a unit of measure, standardized values may be defined as

$$z_{ij} = \frac{x_{ij}}{S(X_i)} . \tag{2-4}$$

The set of all standardized observations will be denoted by the vector, z_i . Thus, each z_i has mean zero and standard deviation equal to unity.

The covariance of z_i and z_k is defined by

$$Cov\,(z_i, z_k) = \frac{1}{n}\sum_{j=1}^{n} z_{ij} \cdot z_{kj} \tag{2-5}$$

and the correlation of z_i and z_k is defined by

$$r_{ik} = \frac{Cov\,(z_i, z_k)}{S(z_i)\,S(z_k)} = \frac{1}{n}\sum_{j=1}^{n} z_{ij} \cdot z_{kj} . \tag{2-6}$$

Let Z be the matrix of standardized observations with typical element, z_{ij}, $(i=1,\cdots,p;\ j=1,\cdots,n)$, and let R_c be the complete matrix of correlation coefficients as defined in equation 2-6. Then,

$$R_c = \frac{1}{n}\, Z Z^{T} \tag{2-7}$$

$$
\frac{1}{n}\underbrace{\begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \end{bmatrix}}_{Z}
\;
\frac{1}{n}\underbrace{\begin{bmatrix} z_{11} & z_{21} & z_{31} \\ z_{12} & z_{22} & z_{32} \\ z_{13} & z_{23} & z_{33} \\ z_{14} & z_{24} & z_{34} \end{bmatrix}}_{Z^T} =
$$

$$
R_c
$$

$$
\frac{1}{n^2}\begin{bmatrix}
z_{11}^2 + z_{12}^2 + z_{13}^2 + z_{14}^2 & z_{21}z_{11} + z_{22}z_{12} + z_{23}z_{13} + z_{24}z_{14} & z_{31}z_{11} + z_{32}z_{12} + z_{33}z_{13} + z_{34}z_{14} \\
z_{11}z_{21} + z_{12}z_{22} + z_{13}z_{23}\ z_{14}z_{24} & z_{21}^2 + z_{22}^2 + z_{23}^2 + z_{24}^2 & z_{31}z_{21} + z_{32}z_{22} + z_{33}z_{23} + z_{34}z_{24} \\
z_{11}z_{31} + z_{12}z_{32} + z_{13}z_{33} + z_{14}z_{34} & z_{21}z_{31} + z_{22}z_{32} + z_{23}z_{33} + z_{24}z_{34} & z_{31}^2 + z_{32}^2 + z_{33}^2 + z_{34}^2
\end{bmatrix} =
$$

$$
R_c^T
$$

$$
\frac{1}{n^2}\begin{bmatrix}
z_{11}^2 + z_{12}^2 + z_{13}^2 + z_{14}^2 & z_{11}z_{21} + z_{12}z_{22} + z_{13}z_{23} + z_{14}z_{24} & z_{11}z_{31} + z_{12}z_{32} + z_{13}z_{33} + z_{14}z_{34} \\
z_{21}z_{11} + z_{22}z_{12} + z_{23}z_{13} + z_{24}z_{14} & z_{21}^2 + z_{22}^2 + z_{23}^2 + z_{24}^2 & z_{21}z_{31} + z_{22}z_{32} + z_{23}z_{33} + z_{24}z_{34} \\
z_{31}z_{11} + z_{32}z_{12} + z_{33}z_{13} + z_{34}z_{14} & z_{31}z_{21} + z_{32}z_{22} + z_{33}z_{23} + z_{34}z_{24} & z_{31}^2 + z_{32}^2 + z_{33}^2 + z_{34}^2
\end{bmatrix} =
$$

$$
\frac{1}{n}\begin{bmatrix} z_{11} & z_{21} & z_{31} \\ z_{12} & z_{22} & z_{32} \\ z_{13} & z_{23} & z_{33} \\ z_{14} & z_{24} & z_{34} \end{bmatrix}
\;
\frac{1}{n}\begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \end{bmatrix}
$$

Figure 1. Formation of the Complete Correlation Matrix

In figure 1, note that corresponding elements of Rc and its transpose are equal and that Rc is, therefore, a symmetric matrix. Further, 12 elements have been combined to form a matrix of 9 elements. In general, there are np elements which combine to form p^2 elements, and, of the p^2 elements, there are $\frac{1}{2}p(p-1)$ distinct elements apart from those on the main diagonal.

The diagonal elements may be determined to be the squares of the sample standard deviations as defined above and are equal to unity. The off diagonal elements are the correlation coefficients defined in equation 2-6.

The correlation matrix, Rc, may be written in the following form which conveys the same information as the representation in figure 1 and is more easily handled:

$$Rc = \begin{matrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{matrix}$$

Figure 2. The Complete Correlation Matrix

CHAPTER 3

THE BASIC MATHEMATICAL MODEL

The basic concept of factor analysis originates in the following three questions:[5]

(1) Can we find some new variables having a linear relationship with the original set of p variables, but fewer in number which account for the observed variation?

(2) If we cannot accomplish (1), is it possible to account for most of the observed variation (to some acceptable degree) with fewer than p variables?

(3) If either (1) or (2) can be accomplished, what are the new variables?

Thus the problem of factor analysis is the problem of determining of a small set of variables ("causes") which adequately account for the variation observed in the original set of variables.

Let z_{ij}, $(i = 1, \cdots, p, \; j = 1, \cdots, n)$, represent a set of n standardized observations on each of p variables such that:

$$\bar{z}_i = \frac{1}{n} \sum_{j=1}^{n} z_{ij} = 0 \qquad 3\text{-}1$$

and

$$S^2(z_i) = \frac{1}{n} \sum_{j=1}^{n} z_{ij}^2 = 1. \qquad 3\text{-}2$$

Let Z be the matrix of standardized observations as described in Chapter 2. Let R_c be the complete correlation matrix defined by

$$R_c = \frac{1}{n} Z Z^T. \qquad 3\text{-}3$$

The following matrix notation will be used to indicate the conformability of the various matrices necessary to formulate the model.

Matrix of Standardized Observations: ${}_pZ_n$ (p rows and n columns).

Matrix of Factor Loadings: ${}_pF_k$ (p rows and k columns),

where k is the number of "causes" present in the individuals that "account" for the deviation.

This matrix specifies the relative importance of the various "causes" in producing the observed "effects." In particular, the element of row i and column j indicates the relative importance of "cause" j on the production of "effect" i.

Matrix of Factor Measurements; ${}_kP_n$ (k rows and n columns).

This matrix specifies the relative strength of the various "causes" as they appear in the individuals. In particular, the element in row j and column s indicates the relative strength of "cause" j as it appears in individual s.

The matrix of factor loadings is made up of three parts: the common factors; the specific factors; and the error factors. Common factors are defined to be factors ("causes") which are applicable (contribute) to more than one of the observed variables. Specific factors are those factors which are peculiar to a single variable. Error factors are applicable to a single factor also.

Notation:

a_{ij} is a factor common to both the ith and jth variables.

s_i is a factor peculiar to the ith variable.

e_i is an error relative to the ith variable.

u_i is a factor unique to the ith variable and is equal to the sum of s_i and e_i.

The elements of a matrix of factor loadings linearly relate the set of new variables with the original set of variables. A matrix of factor loadings is defined as a matrix, not necessarily unique, which, when multiplied by its transpose, is equal to the complete correlation matrix.

$$_pF_{c\,k}F_c^T{}_p = {}_pRc_p \qquad 3\text{-}4$$

The matrix of factor measurements is the set of inferred measurements which have a linear relationship with the original variables, but are fewer in number. It is the mathematical objective of factor analysis to determine in a meaningful sense, the matrix of factor measurements as well as k, the number of new factors.(4)

Figure (3) is presented as an aid in clarifying the common, specific and error factors. In the figure, note that the specific and error factors are involved with the elements along the main diagonal of Rc only. A comparison of figures (2) and (3) will reveal that the elements which are off the main diagonal are the correlation coefficients of figure (2), and that the elements on the main diagonal of figure (3) are equal to the standard deviations of the variables and, as such, are equal to unity. Then, equating the elements of the main diagonals of Rc as represented in figure (2) with corresponding elements of figure (3), we obtain the following relations:

$$\begin{aligned} 1 &= a_{11}^2 + a_{12}^2 + a_{13}^2 + S_1^2 + e_1^2 \\ 1 &= a_{21}^2 + a_{22}^2 + a_{23}^2 + S_2^2 + e_2^2 \\ 1 &= a_{31}^2 + a_{32}^2 + a_{33}^2 + S_3^2 + e_3^2 \end{aligned} \qquad 3\text{-}5$$

Rewriting equations (3-5) so that the right side of the equations consist of the common factors only

$$\begin{aligned} 1 - S_1^2 - e_1^2 &= a_{11}^2 + a_{12}^2 + a_{13}^2 \\ 1 - S_2^2 - e_2^2 &= a_{21}^2 + a_{22}^2 + a_{23}^2 \\ 1 - S_3^2 - e_3^2 &= a_{31}^2 + a_{32}^2 + a_{33}^2 \end{aligned} \qquad 3\text{-}6$$

We may define now a communality to be that part of the variance of a variable which is attributable to the interaction of common factors only. The mathematical definition of a communality, h_i^2, is

$$h_i^2 = \sum_{j=1}^{n} a_{ij}^2, \qquad 3\text{-}7$$

$$
\overset{F}{\begin{bmatrix} a_{11} & a_{12} & a_{13} & s_1 & & & e_1 & & \\ a_{21} & a_{22} & a_{23} & & s_2 & & & e_2 & \\ a_{31} & a_{32} & a_{33} & & & s_3 & & & e_3 \end{bmatrix}}
\overset{F^T}{\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ s_1 & & \\ & s_2 & \\ & & s_3 \\ e_1 & & \\ & e_2 & \\ & & e_3 \end{bmatrix}} = R_c
$$

$$
R_c = \begin{bmatrix} a_{11}^2 + a_{12}^2 + a_{13}^2 + s_1^2 + e_1^2 & a_{21}a_{11} + a_{22}a_{12} + a_{23}a_{13} & a_{31}a_{11} + a_{32}a_{12} + a_{33}a_{13} \\ a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} & a_{21}^2 + a_{22}^2 + a_{23}^2 + s_2^2 + e_2^2 & a_{31}a_{21} + a_{32}a_{22} + a_{33}a_{23} \\ a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} & a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} & a_{31}^2 + a_{32}^2 + a_{33}^2 + s_3^2 + e_3^2 \end{bmatrix}
$$

Figure 3. Common, Specific and Error Factors used to form the Complete Correlation Matrix

or $$h_i^2 = 1 - s_i^2 - e_i^2 = 1 - u_i^2. \qquad 3\text{-}8$$

The concept of communalities will be discussed in Chapter 4 of this paper.

If the specific and error factors are removed from the factor loading matrix, the reduced factor loading matrix, F, remains. The product of the reduced factor loading matrix and its transpose is defined as the reduced correlation matrix, R. The reduced correlation matrix is identical to the complete correlation matrix except that the entries on the main diagonal of the reduced correlation matrix are less than unity. In fact, the diagonal entries of the reduced correlation matrix are the communalities, h_i^2.

$${}_pR_p = {}_pF_r\,F_p^T \qquad 3\text{-}9$$

The reduced correlation matrix is of the following form:

$$R = \begin{bmatrix} h_1^2 & r_{12} & \cdots & r_{1p} \\ r_{21} & h_2^2 & \cdots & r_{2p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ r_{p1} & r_{p2} & \cdots & h_p^2 \end{bmatrix}$$

Figure 4. The Reduced Correlation Matrix

A second approach to the basic model is presented as a means of correlating two primary attacks on the problem and to demonstrate that, in fact, they lead to the identical model.

Factor analysis postulates an underlying structure of a set of measurements in terms of hypothetical variables depending on common and specific factors.[6]

Let $z_{ij} = a_{ij} + u_{i}, (i=1,\cdots,p; j=1,\cdots,n)$ represent n observations on each of p variables, where z_i, a row vector of n observations on the ith variable, has mean zero and variance unity. Then, $z_i = a_i + u_i$, where $u_i = s_i + e_i$ as previously defined, subject to the following restrictions: 3-10

$$Cov(a_i, u_j) = Cov(a_j, u_i) = Cov(u_i, u_j) = 0, \quad i \neq j. \tag{3-11}$$

Then,

$$Cov(z_i, z_j) = Cov(a_i, a_j) + Cov(a_i, u_j) + Cov(u_i, u_j). \tag{3-12}$$

But, applying equation 3-11 yields

$$Cov(z_i, z_j) = Cov(a_i, a_j). \tag{3-13}$$

In matrix notation, let Z, Γ, and Δ be the dispersion matrices (variances and covariances) of the vector variables z_i, a_i, and u_i, then

$$_pZ_n = {_p\Gamma_n} + {_p\Delta_n} \tag{3-14}$$

We can observe the values of the elements of the matrix Z, but not the elements of Γ and Δ, the existence of which is postulated.(6) Factor analysis is mainly concerned with the estimation of the matrix of common factors, starting with Z, the matrix of standardized observations. Equation 3-13 may be written as follows:

$$_pZ_n Z_p^T = {_p\left[{_p\Gamma_n} + {_p\Delta_n}\right]_n} \left[{_p\Gamma_n} + {_p\Delta_n}\right]_p^T \tag{3-15}$$

For clarity, the right member of equation 3-15 is written in expanded form in figure (5).

Expanding the terms of figure (5) and applying equation 3-11 yields the expression in figure (6) since all terms involving the products a_{ij} times u_i are all zero, for they represent $Cov(a_{ij}, u_i), \; i \neq j$.

In figure (6), the off-diagonal elements represent the correlation coefficients, and the main diagonal elements are the variance terms

$$
\left[\Gamma + \Delta\right] \left[\Gamma + \Delta\right]^T
$$

$$
\begin{bmatrix} a_{11}+U_1 & a_{12} & a_{13} \\ a_{21} & a_{22}+U_2 & a_{23} \\ a_{31} & a_{32} & a_{33}+U_3 \end{bmatrix}
\begin{bmatrix} a_{11}+U_1 & a_{21} & a_{31} \\ a_{12} & a_{22}+U_2 & a_{32} \\ a_{13} & a_{23} & a_{33}+U_3 \end{bmatrix} =
$$

$$
\begin{bmatrix}
(a_{11}+U_1)^2 + a_{12}^2 + a_{13}^2 & a_{21}(a_{11}+U_1) + a_{12}(a_{22}+U_2) + a_{13}a_{23} & a_{31}(a_{11}+U_1) + a_{32}a_{12} + a_{13}(a_{33}+U_3) \\
a_{21}(a_{11}+U_1) + a_{12}(a_{22}+U_2) + a_{13}a_{23} & a_{21}^2 + (a_{22}+U_2)^2 + a_{23}^2 & a_{31}a_{21} + a_{32}(a_{22}+U_2) + a_{23}(a_{33}+U_3) \\
a_{31}(a_{11}+U_1) + a_{32}a_{12} + a_{13}(a_{33}+U_3) & a_{31}a_{21} + a_{32}(a_{22}+U_2) + a_{23}(a_{33}+U_3) & a_{31}^2 + a_{32}^2 + (a_{33}+U_3)^2
\end{bmatrix}
$$

Figure 5. Theoretical Method of Forming the Complete Correlation Matrix

$$
\begin{bmatrix}
a_{11}^2 + a_{12}^2 + a_{13}^2 + U_1^2 & a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} & a_{31}a_{11} + a_{32}a_{12} + a_{33}a_{13} \\
a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} & a_{21}^2 + a_{22}^2 + a_{23}^2 + U_2^2 & a_{31}a_{21} + a_{32}a_{22} + a_{33}a_{23} \\
a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} & a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} & a_{31}^2 + a_{32}^2 + a_{33}^2 + U_3^2
\end{bmatrix}
$$

Figure 6. The Complete Correlation Matrix

and are equal to unity. This matrix, therefore, is equivalent element by element to the matrix Rc as defined in equation 3-3 and, thus, the two approaches are identical.

The basic model may be stated as follows:

$$_{p}Z_{n} = {_{p}F_{k}}\,{_{k}P_{n}}. \qquad 3\text{-}16$$

If k is equal to p, then F is square and we can multiply both sides of equation 3-16

$$_{k}F^{-1}_{p}\,{_{p}Z_{n}} = {_{k}F^{-1}_{p}}\,{_{p}F_{k}}\,{_{k}P_{n}},$$

$$_{k}F^{-1}_{p}\,{_{p}Z_{n}} = {_{k}I_{k}}\,{_{k}P_{n}}, \qquad 3\text{-}17$$

or $$F^{-1}Z = P.$$

Since we are attempting to establish the meaningful determination of k and the matrix of factor measurements, our problem is simple when k is equal to p, but in all probability, we have accomplished very little of value, since we have the same number of variables as before.

If k is less than p, we can operate as follows:

$$_{p}R_{p}\,R^{-1}_{p}\,Z_{n} = {_{p}F_{k}}\,P_{n} \qquad 3\text{-}18$$

or $$R\,R^{-1}Z = F\,P. \qquad 3\text{-}19$$

Since FF^T is equal to R, we can substitute FF^T in equation 3-19

$$F\left(F^{T}R^{-1}Z\right) = FP \qquad 3\text{-}20$$

and $F^TR^{-1}Z$ is equivalent to P as a post-multiplier of F.

We will define now the matrix of factor weights to be

$$W \equiv F^{T}R^{-1} \qquad 3\text{-}21$$

since it is the linear operator by means of which P is expressed in terms of Z. If k is equal to p

$$F^{T}R^{-1} = F^{-1} \qquad 3\text{-}22$$

and $$_{k}P_{n} = {_{k}W_{p}}\,{_{p}Z_{n}}. \qquad 3\text{-}23$$

The process of standardizing the set of observations on each of the variables in effect makes a set of vectors which are all of the same

length so that each has an equal effect on the outcome of factoring or rotation. In many problems, it is not desirable to permit each vector to exercise equal influence. For example, in an article relative to patterns of magazine reading by the public, co-variances were used because it seemed appropriate to weigh each relationship in terms of the number of persons involved. If correlations had been used, magazines with small homogeneous audiences would have tended to dominate the analysis.[1]

---

[1]Twery, R. J.,"Detecting Patterns of Magazine Reading," J. of Marketing, January 1958, pp. 290-294.

CHAPTER 4

COMMUNALITIES AND RANK

The primary difficulty in factor analysis is the selection of "good" values for the communalities. The rank of the correlation matrix and, accordingly, the number of factors is dependent on the selection of the communalities.

Values for the communalities may range from very close to zero up to a maximum value of one. Theproblem of selection of the appropriate communality may be restated in terms of the purpose of factoring.

In general, one will engage in a factor analysis for one of two reasons:

1. Factoring to determine a hypothesis, to predict, or simply looking to see if there is some underlying order to the variables joined in the correlation matrix;

2. Factoring, where the factoring is the end product. That is the determination of relatively precise values of the factor loadings. In this case, a good estimate of the communalities is essential. The inexperienced factorist will have to read widely on the subject of communalities in order to develop perspective and skill in factoring to the point where he may be considered an artisan, or refer the problem to an expert in the field of factor analysis.

If, however, the purpose of the factoring is as stated in paragraph 1 above, then there are three possible approaches to the choice of the communalities:

1. Use upper bounds for the communalities which tend to overestimate the rank of the correlation matrix and the individual factor loadings;

2. Use lower bounds for the communalities which leads to a conservative approach with underestimation of the rank and factor loadings; or

3. Use one of the many rule of thumb estimates for the communalities and hope that the result is not too far off.

The upper bounds for the communalities are unities for a correlation matrix and the appropriate variance for a covariance matrix. The use of the upper bounds along the main diagonal of the matrix implies that all variance is caused by the common factors alone. Obviously, this is not a correct picture, but it may be useful when probing a matrix to see what can be done with it.

The lower bounds for the communalities may be determined by the use of squared multiple correlations along the main diagonal. Guttman has shown that the squared multiple correlation of any variable with all other variables is a lower bound for the communality of that variable.[1]

Rules for subjectively determining the communalities are to be found on pages 282 to 318 of Multiple Factor Analysis by L. L. Thurstone,[(8)] and in most of the general books on the subject of factor analysis. Such rules must be used with caution, for many of the techniques of factoring are very sensitive to the selected values of the communalities.

There is an iterative procedure to compute communalities. Factor the correlation matrix with initial estimates of the communalities. Multiply the factor loading matrix produced by its transpose to produce a correlation matrix. Insert the diagonal entries of the computed correlation matrix into corresponding positions along the main diagonal of

---

[1]Guttman, L. "Multiple Rectilinear Prediction and Resolution into Components," Psychometrika, 1940, 5, 75-99.

the original correlation matrix. Repat the process until the communalities stabilize on a value. Although it might seem that this procedure should produce a very good estimate for each communality, unfortunately there is no assurance that the values will ever converge. In fact, it has been shown by actual experiment where an iteration was run through over 100 factorings that the communalities did not necessarily stabilize. Therefore, regardless of the appeal of the iteration method for the determination of the communalities, this technique must be rejected.[2]

When factoring is done on a computer such as the CDC 1604 at the U. S. Naval Postgraduate School, it should be feasible to factor each correlation matrix at least three times; once using upper bounds; once with estimated values for the communalities; and once with squared multiple correlations used as the communalities. That is, factor over the extremes. The results thus produced may give insight into the underlying structure which otherwise might not have been available to the factorist.

Since the rank of the correlation matrix determines the number of factors which may be extracted, a short discussion relative to the rank of the matrix is appropriate to a discussion of the number of factors and communalities.

The rank of the reduced correlation matrix can always be reduced to k, where k is the smallest integer greater than
and p is the order of the matrix.[3]

---

[2]Twery, R. J. Interview with the author.

[3]Lederman, W. "On the Rank of the Reduced Correlation Matrix in Multiple Factor Analysis," Psychometrika, 1937, 2, 85-93.

Dr. C. F. Wrigley and Mr. R. J. Twery have developed a method for establishing a lower limit for the number of factors. Mr. Twery has kindly consented to the inclusion of that method in this paper.

The rank of any (correlation) matrix is equal to the number of positive latent roots of the matrix. Hence, the determination of a lower limit for the number of positive latent roots of the reduced correlation matrix, R, is equivalent to the establishment of a lower limit for the rank, and, accordingly, a lower limit for the number of common factors. If the factor analysis mathematical model holds exactly, then the number of common factors is the same as the rank of the correlation matrix.

Guttman has shown that the squared multiple correlation of any variable with all other variables is a lower bound for the communaltiy of the variable. Let $\bar{R}$ be the correlation matrix with those lower bounds for the communalities inserted in the main diagonal. We shall prove a general lemma which is sufficient to show that the number of positive latent roots of $\bar{R}$ is a lower bound for the number of positive roots of R. We shall then have shown that a lower bound exists for the number of common factors. Lemma: The latent roots of any symmetric matrix do not decrease when a positive increment is made to any entry in the main diagonal.

Since $(R-\bar{R})$ is a diagonal matrix, R can be expressed as the sum of $\bar{R}$ and p matrices which are null except for one diagonal element in each. The latent roots of the p partial sums which are successively formed in the process of building R from the p 1 terms must, by the lemma, be monotone non-decreasing. It follows that there will be at least as many positive latent roots of R as of $\bar{R}$.

Computational procedure.

1. Calculate the squared multiple correlation of each variable with the other p-1 variables involved in the matrix.

2. Calculate the latent roots and latent vectors for the correlation matrix with the squared multiple correlations.

Let there be found to be m positive roots. By the proof given above, we know that there are at least m common factors. It should be noted that these calculations are quite practicable on a computer.

It is proposed that this procedure should be used to decide upon the number of factors to be included in any rotation of the initial results.

In summary.

If the rank of the correlation matrix is large and large factor loadings are anticipated, then be generous, be conservative, use the squared multiple correlations which are the lower bounds for the communalities and be sure of what is produced. Certain extra advantages accrue to a factorist employing this method, for the method is explicit and objective. When using a computer, a program for determining the inverse of a matrix is almost always available. That affords a straightforward technique for computing the squared multiple correlations.

If it is anticipated that the rank of the correlation matrix will be small and that few factors will be extracted, go to the other extreme and overestimate all the way. Use unities in the main diagonal of the correlation matrix and squeeze out all the factors possible.

BIBLIOGRAPHY - CHAPTER 4

Albert, A. A., The Matrices of Factor Analysis, Proc. Nat. Acad. Sci., 1944, 30, pp. 90-95.

The Minimum Rank of a Correlation Matrix, Proc. Nat. Acad. Sci., 1944, 30, pp. 144-146.

Burt, C., Alternative Methods of Factor Analysis and their Relations to Pearson's Method of Principal Axes, Brit. J. Psych., Stat. Sec., 1949, 2, pp. 98-121.

The Factorial Study of Temperamental Traits, Brit. J. Psych., Stat. Sec., 1948, 1. pp. 178-203.

Tests of Significance in Factor Analysis, Brit. J. Psych., Stat. Sec., 1952, 5, pp. 109-133.

Coombs, C. H., A Criterion for the Significance of Common Factor Variance, Psychometrika, 1941, 6, pp. 267-272.

Davis, F. B., Fundamental Factors of Comprehension in Reading, Psychometrika, 1944, 9, pp. 185-197.

Eckart, C. and Young, G., The Approximation of One Matrix by Another of Lower Rank., Psychometrika, 1936, 1, pp. 211-218.

Gullikson, H., Theory of Mental Tests, Wiley, 1950.

Guttman, L., Multiple Rectilinear Prediction and Resolution into Components, Psychometrika, 1940, 5, pp. 75-100.

A New Approach to Factor Analysis: The Radex, in P. F. LAZARFELD Ed., Mathematical Thinking in the Social Services. The Free Press, 1954.

Some Necessary Conditions for Common Factor Analysis, Psychometrika, 1956, 21, pp. 273-285.

Guttman, L., Improved Bounds for Communalities, AF Contract 41(751)-76, Report 6, Univ. of Cal., 1956.

To What Extent Can Communalities Reduce Rank?, AF Contract 41(657)-76, Report 8, Univ. of Cal., 1956.

Image Theory for the Structure of Quantitative Variates, Psychometrika, 1953, 18, pp. 277-296.

Heywood, H. B., On Finite Sequences of Real Numbers, Proc. Roy. Soc., Sec. A., 1931, 134, pp. 486-501.

Hotelling, H., Analysis of a Complex of Statistical Variables into Principal Components, J. Educational Psych., 1933, pp. 417-441 and pp. 498-520.

Householder, A. S. and Young, G., Matrix Approximation and Latent Roots, Amer. Math. Monthly, 1938, 45, pp. 165-171.

Kaiser, H. F., Solution for the Communalities: A Preliminary Report, AF Contract 41(657)-76, Report 5, Univ. of Cal., 1956.

Lawley, D. N., The Estimation of Factor Loadings by the Method of Maximum Likelihood, Proc. R. Soc. Edin., 1940, 60, pp. 64-82.

Lederman, W., On the Rank of the Reduced Correlation Matrix in Multiple Factor Analysis, Psychometrika, 1937, 2, pp. 85-93.

On a Problem Concerning Matrices with Variable Diagonal Elements, Proc. R. Soc., Edin., 1939, 60, pp. 1-17.

McNemor, O., On the Number of Factors, Psychometrika, 1942, 7, pp. 9-18.

Rippe, D. D., Application of Large Sampling Criterion to Some Sampling Problems in Factor Analysis, Psychometrika, 1953, 18, pp. 191-205.

Roff, M., Some Properties of the Communality in Multiple Factor Theory, Psychometrika, 1936, 1 (June), pp. 1-6.

Rosher, B., An Algebraic Solution for the Communalities, Psychometrika, 1948, 13, pp. 181-184.

Thurstone, L. L., Note on a Re-Analysis of Davis' Reading Tests, Psychometrika, 1946, 11, pp. 185-188.

Tryon, R. C., Cluster Analysis, Edwards Bros., 1939.

Communality of a Variable: Formulation by Cluster Analysis, AF Contract 41(657)-76, Report 4, Univ. of Cal., 1956.

Wherry, R. J., A New Iterative Method for Correcting Erroneous Estimates in Factor Analysis, Psychometrika, 1949, 14, pp. 231-234.

Wrigley, C. F., Some Observations on the Distinction Between Common and Specific Variance in Factor Analytic Theory, AF Contract 41(657)-76, Report 2, Univ. of Cal., 1956.

Wrigley, C. F. and Twery, R. and Golub, G. H., The Distinction Between Common Factors and Specific Factors, Amer. Psychologist, 1954, 9, pp. 496-497, Abstract.

Young, G., Matrix Approximation and Subspace Fitting, Psychometrika, 1937, 2, pp. 21-26.

Factor Analysis and the Index of Clustering, Psychometrika, 1939, 4, pp. 201-208.

*Ezekiel, M., Methods of Correlation Analysis, Wiley, 1941.

*DuBois, P. H., Multivariate Correlation Analysis, Harper, 1957.

*Multiple Correlations

CHAPTER 5

THE SHORTENED SQUARE ROOT METHOD OF FACTOR ANALYSIS

The square root or diagonal method of factoring was developed by L. L. Thurstone.(9) As developed by Thurstone, it has proven to be very sensitive to "good" values of the estimated communalities.

Before attempting to describe the shortened square root method of factoring, an outline of the original method will be furnished.

The diagonal method consists in expressing the correlation matrix, R, in terms of the product of F and its transpose where F is a triangular matrix.(3)

		FACTOR				
		I	II	III	IV	V
Variables	1	a_{11}	0	0	0	0
	2	a_{21}	a_{22}	0	0	0
	3	a_{31}	a_{32}	a_{33}	0	0
	4	a_{41}	a_{42}	a_{43}	a_{44}	0
	5	a_{51}	a_{52}	a_{53}	a_{54}	a_{55}

Figure 7. A Triangular Factor Loading Matrix

The procedure to accomplish the triangularization of R is straightforward. Calculate loadings on the first factor which reduce the correlation for a selected "pivot" variable to zero. Compute residual correlations and factor loadings on the second pivot variable so that the residual correlations for the second variable are also reduced to zero. The process continues until the diagonal factor loading matrix is produced. Successive reference axes are orthogonal so that the factor

measurements are uncorrelated. When all p factors have been extracted, the residual correlations are all zero.

The general computational procedure is as follows:
Let a_{i1} be the factor loading for the ith variable on the first factor, then, since $FF^T = R_1$

$$a_{i1} = \frac{r_{ip}}{\sqrt{r_{pp}}} \qquad 5\text{-}1$$

where r_{ip} is the correlation between the ith variable and the pivot variable. Note that in this case, p refers to pivot variable which may be any of the variables represented in the correlation matrix.

The column and row of the pivot variable in the residual correlation matrix will consist of zeros for

$$r_{1,ip} = r_{ip} - \frac{r_{ip}}{\sqrt{r_{pp}}}\frac{r_{pp}}{\sqrt{r_{pp}}} = 0, \ (i = 1, \cdots, p). \qquad 5\text{-}2$$

Similarly, the column and row of each subsequent pivot variable in the appropriate residual correlation matrix will consist of zeros. Therefore, after extracting p factors, the residual correlations will all be zero.

If the rank of R is k, the kth residual correlation matrix will consist of all zeros and the loadings on the k ist and all subsequent factors will be zero.

Factoring in the shortened square root method[1] commences with Rc, the complete correlation matrix, (i.e., unities along the main diagonal of the correlation matrix). Each column of Rc is added omitting the elements along the main diagonal.

$$\sum_{i=1}^{p} r_{ij}, \ (j = 1, \cdots, p; \ i \neq j) \qquad 5\text{-}3$$

---

[1]Wrigley, C. and McQuitty, L.L. The Square Root Method of Factor Analysis: A re-examination and a shortened procedure, AF Contract 33(038)-25726, Memo, Report A-9, 18 August 1959.

The first pivot variable is selected as that variable having the largest absolute value of the column sums. If the largest sum is negative, the signs will be reversed for the appropriate factor in the factor loading matrix. Once the pivot variable has been selected, the first factor loading may be computed as follows:

$$a_{j1} = \frac{r_{jp}}{\sqrt{r_{pp}}} . \qquad 5\text{-}4$$

But r_{pp} is equal to unity on the first factor loading only. Accordingly, the first factor loadings are simply the elements in the column of the pivot variable in the correlation matrix. Note that this is true for the first factor only. That is to say

$$a_{j1} = r_{jp} \quad (j = 1, \cdots, p). \qquad 5\text{-}5$$

Once the factor loadings have been determined, the column sums of the residual correlation matrix are calculated for all variables. Residual correlations are defined to be that part of the original correlations which are left to be accounted for after one or more factors have been extracted. (Residual correlations are added algebraically and main diagonal entries are omitted from the column sum.) The second pivot variable is selected using the same criteria as before. That is, it is the variable which has the largest absolute value of the column sums. The column sums of the residual correlations are calculated by use of:

$$\sum_{i=1}^{p} r_{k,ij} = \sum_{i=1}^{p} r_{k-1,ij} - a_{jk} \sum_{i=1}^{p} a_{ik} , \quad i \neq j \qquad 5\text{-}6$$

where $r_{k,ij}$ is the kth residual correlation and $r_{,-1,ij}$ is the k-1st residual correlation between the ith and the jth variables.

Once the pivot variable has been determined, the residuals are calculated for that variable, say the sth, as follows:

$$r_{k+1,is} = r_{is} - \sum_{t=1}^{k} (a_{it} \cdot a_{st}) \qquad 5\text{-}7$$

where a_{it} and a_{jt} are loadings for the ith and sth variables already extracted. $(t = 1, \ldots, k)$

The factor loadings for the next factor are

$$a_{i,k+1} = \frac{r_{k,is}}{\sqrt{r_{k,ss}}} \qquad 5\text{-}8$$

This process is repeated until all of the factors have been extracted so that a diagonal matrix of factor loadings has been produced.

Appendix I contains a simple numerical example of this method which may prove helpful in interpreting the foregoing discussion.

There are several advantages to the shortened square root method of factoring.[2]

1. Each step is simple both conceptually and computationally.

2. Considerable time is saved over other methods for there are fewer calculations. This saving can be reflected in the ease of preparing a program for a computer as well as the saving in memory space.

3. A computer program can be arranged so that there will be no human action required once the program and values have been loaded into the computer. The computer can be instructed to calculate the correlations and the square root factor loadings for any desired number of factors. The method is capable of factoring very large matrices efficiently.

4. When factoring by hand or with a desk calculator, checks are available at every important step in the procedure. See Appendix I for a demonstration of the check procedures.

---

[2] Wrigley, C. and McQuitty, L.L., The Square Root Method of Factor Analysis: A Re-examination and a Shortened Procedure, Memo Report A-9, AF Contract 33(038)-25726, August 18, 1956.

5. The square root method as described herein has been found to give more clear results than either the centroid or the principal axes methods of factoring. This statement may be justified as follows: there are many more positive than negative factor loadings due to the selection of pivot variables with high sums of residual correlations; also, there are $k(k-1)/2$ zeros (where k is the number of factors extracted) in the factor loading matrix; and any square root factor loading is determined by the correlation or residual correlation of the ith and kth variables only, while other methods require averaging over a set of the k-1 residual correlations with the ith variable. Accordingly, the residual correlation with a single variable is generally more easily interpreted than an average over a set of variables.

Although this method has several distinct advantages, it necessarily has certain disadvantages:

1. The loading for the pivot variable is always much greater than that for any of the other variables. This may result in the pivot variable being favored in any subsequent rotation.

2. There is the universal problem of deciding when to stop factoring. This problem must be resolved regardless of the method of factoring employed. When working with computers, the analyst may deliberately overfactor as a means of insuring that he has extracted all meaningful factors. The factors may then be arranged in order of their contribution to the variance. Those which contribute little or nothing may be discarded. It is worth noting that the order of factoring in this method has nothing to do with the relative contributions of the various factors. The last factor extracted may well be the most critical factor in the whole matrix.

CHAPTER 6

THE CENTROID METHOD OF FACTORING

The complete centroid method of factoring has been widely employed to factor correlation matrices where experimental data are involved.[8]

The method does not produce a mathematically unique solution.

The centroid method can best be described in terms of geometry, for it utilizes the centroid concept much as in physics. The variables may be considered as a set of p vectors which are contained in a space of k dimensions, where k is the number of common factors, or k is the dimension of the common factor space. The "dot" product between any pair of vectors is the correlation between the components of these vectors.

A numerical example is provided in Appendix II so that the development may be followed more easily.

Since the correlations are independent of the coordinate axes, the reference axes may be rotated so that the centroid lies on one of the axes, say the first.

Starting with a normal factor pattern, the correlations are computed as outlined in Chapter II. Then, making the assumption that the residual correlations vanish, that is to say that the extracted factors completely account for all of the correlation. The observed correlations may be written

$$r_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{ik}a_{jk} \qquad 6\text{-}1$$

Thurstone refers to this equation as the fundamental relation of factor analysis.[8]

The notation, a_{i1}, denotes the first centroid factor loading relative to the ith variable. It may be interpreted, also, as the ith coordinate of the vector .

The centroid has coordinates

$$\frac{1}{p}\sum_{i=1}^{p} a_{i1},\ \frac{1}{p}\sum_{i=1}^{p} a_{i2},\ \ldots,\ \frac{1}{p}\sum_{i=1}^{p} a_{ik}. \qquad 6\text{-}2$$

If the centroid is located on the first coordinate axis, then it will have coordinates

$$\frac{1}{p}\sum_{i=1}^{p} a_{i1},\ 0,\ \ldots,\ 0. \qquad 6\text{-}3$$

It is possible to determine the coordinates of the first centroid factor by adding each column of the correlation matrix and the associated factor loading matrix:

$$\sum_{i=1}^{p} r_{ij} = a_{j1}\sum_{i=1}^{p} a_{i1} + a_{j2}\sum_{i=1}^{p} a_{i2} + \cdots + a_{jk}\sum_{i=1}^{p} a_{ik}. \qquad 6\text{-}4$$

Applying equation 6-3 yields:

$$\sum_{i=1}^{p} r_{ij} = a_{j1}\sum_{i=1}^{p} a_{i1}. \qquad 6\text{-}5$$

The sum of all of the elements of the correlation matrix is given by

$$\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij} = \sum_{j=1}^{p} a_{j1}\sum_{i=1}^{p} a_{i1} = \left(\sum_{i=1}^{p} a_{i1}\right)^2. \qquad 6\text{-}6$$

Then,

$$\sum_{i=1}^{p} a_{i1} = \pm\sqrt{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}} \qquad 6\text{-}7$$

We will arbitrarily take the positive square root and substitute in equation 6-5 to obtain the following expression:

$$a_{j1} = \frac{\sum_{i=1}^{p} r_{ij}}{\sqrt{\sum_i\sum_j r_{ij}}} = \frac{t_j}{\sqrt{T}} \qquad 6\text{-}8$$

where t_i is the sum of all of the correlations in column i of the correlation matrix and T is the total of all of the elements in the correlation matrix. The selection of the positive square root is not essential, for, if the negative had been chosen, it may be shown that an equally acceptable solution would have been obtained.[4]

The use of equation 6-8 will produce the coefficient of the first centroid factor, a_{i1}, for the ith variable which may be interpreted as the first coordinate for each point representing each variable. It should be noted that this equation is based on the assumption that the centroid is not at the origin, for if it were, the indicated division would not be possible.

The first residual correlation resulting from the extraction of the first factor must be computed. Residual correlations with one, two, ..., k-1 factors removed are utilized in successive stages of the centroid method. The following notation will be used to designate a residual correlation:

$r_{1,ij} \equiv$ The first residual correlation on the variables Z_i and Z_j.

$r_{k-1,ij} \equiv$ The k-1st residual correlation on the variables.

The residual correlations may be calculated by use of the following equation:

$$\begin{aligned} r_{1,ij} &= r_{ij} - a_{i1}a_{j1} = a_{i2}a_{j2} + \dots + a_{ik}a_{jk} \\ r_{2,ij} &= r_{1,ij} - a_{i2}a_{j2} = a_{i3}a_{j3} + \dots + a_{ik}a_{jk} \\ r_{k-1,ij} &= r_{k-2,ij} - a_{i,k-1}a_{j,k-1} = a_{ik}a_{jk}. \end{aligned} \qquad 6\text{-}9$$

The residual correlations may be regarded as the dot product of pairs of residual vectors in a space of one less dimension than the initial space.

The k-1 coordinates of the centroid are:

$$\frac{1}{p}\sum_{i=1}^{p} a_{i\tau}, \quad (\tau = 1, \dots, k). \qquad 6\text{-}10$$

Reference to equation 6-3 reveals that all of the summations in equation 6-10 are zero and that the centroid is at the origin in k-1 space.

In order to proceed with the centroid method, it will be necessary to move the centroid away from the origin. Reflection in the origin has the effect of measuring the reflected vector in the opposite direction, i.e.:

$$Z_i = - Z_i \text{ REFLECTED.} \qquad 6\text{-}11$$

If we reflect the vector Z_i in the origin and hold the other vectors fixed, the result will be that all of the signs of the correlations between the ith and the other variables will be reversed.

Thurstone states "it is desirable to account for as much as possible of the residual variance by each successive factor."(8) In order to follow Thurstone's principle, it is desirable to move the centroid as far as possible from the origin. This may be accomplished by reflecting those variables which have the greatest number of negative correlations to bring them into the hyperhemisphere where the majority of the variables, which have few or no negative correlations, lie.

In Appendix II, the process of reflecting variables (vectors) is described in detail.

Once the residual correlations have been calculated and the centroid moved away from the origin by reflections, the factoring process is repeated until the desired number of factors has been extracted.

Each successive centroid axis is at right angles to all of the previous centroid axes since the residual correlation space is orthogonal to the space of the centroid axes related to factors which have been extracted. By extracting each centroid factor, the residual correlations are reduced in magnitude. Furthermore, the rank of each successive residual correlation matrix is one less than the rank of its predecessor.

In this development, no restrictions have been imposed on the elements along the main diagonal of the correlation matrix. The procedures and suggestions described in Chapter 4 relative to the communalities, rank and number of factors are applicable to the centroid method of factoring.

CHAPTER 7

OTHER METHODS OF FACTORING

The two methods of factoring described in detail in this paper are by no means representative of the entire field of factoring a correlation matrix. However, they should give the beginner an idea as to how to go about factoring.

Two other methods of importance will be discussed briefly. If additional information is desired regarding these methods, attention is invited to the bibliography at the end of this chapter.

The principal axes method resolves the problem of the indeterminacy of the reference frame by placing the first reference axis so as to maximize the sum of the squares of the first factor loadings.[3] The next axis will be located to maximize the sum of the squares of the second factor loadings. The process is continued such that each successive reference axis is located so that the sum of the squares of the corresponding factor loadings is maximized.

One of the advantages of the principal axes method is that an unique solution is produced. The method is useful to condense the variables by expressing them in terms of a relatively few linearly independent factors.

The square root solution may be converted to a principal axes solution by performing the following operation:

Let F be the square root solution and F_p, the principal axes solution orthogonal by columns, then F_p is equal to FU where U is defined as the matrix of latent vectors of the product matrix .

The maximum likelihood method was developed by Lawley and Rao in slightly different methods.

Observed correlation coefficients are continuous random variables. Although the population matrices may be reducable to small minimum rank,

the sample correlation matrices are not reducable generally to the minimum rank of the population matrix. This situation suggested the following statistical problems: Obtain an estimate, , of the minimum rank of the variance-covariance matrix of the common parts of the variables given by the relations presented in chapter 2. Test the hypothesis $\Omega \leq \hat{k}$, $\Omega = 1, \cdots, n-1$ against the alternative $\Omega > \hat{k}$. It is necessary that a basis for the common factor space be estimated.

Lawley set up a likelihood function for the factor loadings for any number of factors.[1] He then derived a large sample statistical test for the hypothesis, given the maximum likelihood factor loadings.

Rao formulates the problem: "What is that factor variable which is predictable from Z_i with the maximum possible precision"?[2] For any number of factors, Rao provides a test of significance which gives a lower confidence limit for the number of factors. Rao has shown that his resolution of the problem is equivalent to that of Lawley.

The maximum likelihood method presents a solution to the communality problem.

---

[1] Lawley, D. H., The Estimation of Factor Loadings by the Method of Maximum Likelihood, Proc. Roy. Soc., Edin., 1940, 60, pp. 64-82.

[2] Rao, C. R., Estimation and Tests of Significance in Factor Analysis, Psychometrika, 1955, 20, pp. 93-111.

Bartlett, M. S., Internal and External Factor Analysis, Brit. J. Psychology, Stat. Sec., 1948, 1, pp. 73-81.

Tests of Significance in Factor Analysis, Brit. J. Psychology, Stat. Sec., 1950, 3, pp. 77-85.

The Effect of Standardization on a chi square Approximation in Factor Analysis, Biometrika, 1951, 38, pp. 337-444.

Grishick, M. A., Principal Components, J. Amer. Statist Association, 1936, 31, pp. 519-528.

Hoel, P. G., A Significance Test for Component Analysis, Ann. Math. Stat., 1937, 8, pp. 149-158.

A Significance Test for Minimum Rank in Factor Analysis, Psychometrika, 1939, 4, pp. 245-254.

Hotelling, H., Analysis of a Complex of Variables into Principal Components, Psychometrika, 1936, 1, pp. 27-35.

Kelley, T. M., Essential Traits of Mental Life, Harvard Studies in Education, 26, Harvard Univ. Press, 1935, Chap. 1.

Thurstone, L. L., Multiple Factor Analysis, Psychological Review, XXXVIII, 1931, pp. 406-427.

Theory of Multiple Factors, Edwards Bros., Ann Arbor, Mich., 1932, pp. 17-19 and 22-27.

Wilson, E. B. and Worchester, J., Note of Factor Analysis, Psychometrika, 1939, 4, pp. 133-148.

Emmett, W. G., Factor Analysis by Lawley's Method of Maximum Likelihood, Brit. J. Psychology, Stat. Sec., 1949, 2, pp. 90-97. See also correction, Vol. 3, p. 76.

Henrysson, S., The Significance of Factor Loadings: Lawley's Test Examined by Artificial Samples, Brit. J. Psychology, Stat. Sec., 1950, 3, pp. 159-165.

Lawley, D. N., The Estimation of Factor Loadings by the Method of Maximum Likelihood, Proc. Roy. Soc., Edin., 1940, 60, pp. 64-82.

The Application of the Maximum Likelihood Method to Factor Analysis, Brit. J. Psychology, 1943, 33, pp. 172-175.

Further Investigations in Factor Estimation, Proc. Roy. Soc., Edin., 1941, 61, pp. 176-185.

Problems in Factor Analysis, Proc. Roy. Soc., Edin., Sec A, 1949, 62, pp. 394-399.

Rao, C. R., Estimation and Tests of Significance in Factor Analysis, Psychometrika, 1955, 20, pp. 93-111.

Thomson, G. H., The Factorial Analysis of Human Ability, U of London Press, 5th Edition, 1951, Chap. 9.

Young, G., Maximum Likelihood Estimation and Factor Analysis, Psychometrika, 1941, 6, pp. 49-53.

Note: Many of the general references listed in the bibliography at the end of this paper include information relative to principal axes and the maximum likelihood methods.

CHAPTER 8

PREFERRED TYPES OF SOLUTION AND ROTATION

For any matrix of correlations, each of the methods of factoring will result in an arbitrary set of reference axes out of an infinite set of axes. In order to telocate the axes in some position useful for interpretation of the factors and for comparison with other studies, the axes are rotated. The objective of rotation is to obtain meaningful factors that are as invariant as possible from one analysis to another.

In order to provide some basis of determining the factor patterns which are preferred, the following standards have been established.[(4)]

1. The pattern supports assumed composition of the variables. That is the initial assumption that each variable was composed of three components, i.e., common factors, unique factors, and error factors. All forms of the preferred factor pattern must conform to the assumed composition of the variables developed in Chapter 3.

2. Parsimony. The total number of common factors should be as small as possible and the linear description of each variable should be as simple as possible.

3. Uncorrelated factors. The orthogonal solution is preferred from several standpoints.[1]

a. Independence. Theoretically the factors represented by orthogonal axes are independent. The use of correlated factors should be restricted to those cases where it can be shown that independent measures cannot be constructed for a given population.

b. Simplicity. Orthogonal systems are much more easily handled, both computationally and graphically.

---

[1] Cattall, R. B., Parallel Proportional Profiles, Psychometrika 9, 1944, pp. 267-283.

c. Instability of angular separation. The obliqueness of the axes in any one factor analysis is determined to some extent by sampling errors and sample variation. Therefore, the separation of the axes can be expected to vary from sample to sample.

d. Similarity of results. There is little difference in the interpretation of the factors obtained from two different methods except for cases of highly correlated factors. Accordingly, the added theoretical and computational complications of oblique rotations produce little of practical value. A secondary problem is to determine how oblique should the axes be before the axes can be considered as separate no longer and be treated as one axis.

4. Correlated factors. Oblique solutions may be useful[2] when the factors are correlated. Proponents of the oblique solution advance the following arguments.

a. Important information. If the factors produced are correlated, this fact should be taken into account in interpreting the results. Also, second order and higher order analyses can be made of the intercorrelations of the factors.

b. Better fit of the Axes. If the axes are passed exactly through clusters of variable vectors representing the data, a more satisfactory structure can be obtained.

c. Closer duplication of Nature. A number of measures as they appear in nature tend to be related. A common example is the correlation usually assumed between the height and weight of an individual. In general, a tall man is heavier than a short man.

---

[2] Ibid.

5. Relative contribution of factors. Three useful types of relationships between the contricutions of a set of factors are:

a. Decreasing contributions, or the factors contribute successively smaller amounts to the total variance.

b. Level contributions, or each factor contributes an equal or anearly equal amount to the total variance.

c. One large and the remaining factors level contributions.

The general references in the bibliography all contain information relative to the rotation of factors.

APPENDIX I

NUMERICAL EXAMPLE OF THE SHORTENED SQUARE ROOT METHOD

In hand calculation or when using a desk calculator, there is a slight loss in precision if the diagonal entries are included in the column sums. Ease of computation, simple checks for accuracy, and increased speed in calculation more than compensate for the loss in accuracy. In this example, we shall include the entries along the main diagonal in all column sums.

The example will be developed in a step by step manner and will include all calculations at every stage.

Step 1. Form the complete correlation matrix as described in Chapter 2. Figure 8 is a hypothetical complete correlation matrix.

Step 2. Add the columns of the matrix algebraically. Form the sum of the column totals. The results are entered in figure 8.

Step 3. Select as the first pivot variable, that variable with the largest column sum. In the example, variable 3 is the first pivot variable. Since the diagonal entries are all unity, the first factor loadings are equal to the correlations in column 3 of the correlation matrix. Enter the elements of column 3 in figure 8 in the column marked I of figure 9.

Step 4. Using equation (5-6) modified in that i is allowed to be equal to j, calculate the column sums of the first residual correlation matrix.

$$\sum_{i=1}^{5} r_{1,ij} = \sum_{i=1}^{5} r_{ij} - a_{j1} \sum_{i=1}^{5} a_{i1} \; ; \; j = 1, \cdots, 5. \qquad \text{I-1}$$

$$\sum_{i=1}^{5} r_{1,i1} = 2.75 - .54 \times 3.24 = 1.00.$$

$$\sum r_{1,i2} = 3.01 - .71 \times 3.24 = .71.$$
$$\sum r_{1,i3} = 3.24 - 1.00 \times 3.24 = 0.$$
$$\sum r_{1,i4} = 2.86 - .49 \times 3.24 = 1.[illegible]1.$$
$$\sum r_{1,i5} = 2.95 - .50 \times 3.24 = 1.3[illegible].$$

Step 5. Enter the column sums in column R_1 of figure 10.

Step 6. Variable 5 is the second pivot variable since it has the largest column sum. Calculate the residual correlations with variable 5 and the other variables using equation (5-7). Enter the values in the first column of figure 11.

$$r_{k,is} = r_{is} - \sum_{T=1}^{k} a_{iT} a_{sT} \; , \quad i = 1, \cdots, 5; \; s = 5.$$

$$r_{1,15} = .38 - .54 \times .50 = .11.$$ I-2
$$r_{1,25} = .41 - .71 \times .50 = .05.$$
$$r_{1,35} = .50 - 1.00 \times .50 = 0.$$
$$r_{1,45} = .65 - .49 \times .50 = .40.$$
$$r_{1,55} = 1.00 - .50 \times .50 = .75.$$

Check 1. Does the column sum of the first column of figure 11 equal the column sum for the pivot variable in figure 10? Neglect rounding errors. $1.31 \simeq 1.32.$ I-3

Check 2. Does $\sum_i \sum_j r_{ij} - \left(\sum_{i=1}^{5} a_{i1} \right)^2 = \sum_i \sum_j r_{1,ij}$?

$$14.80 - 10.50 = 4.30.$$

Step 7. Calculate $\dfrac{1}{\sqrt{r_{1,88}}} = \dfrac{1}{\sqrt{.75}} = 1.154.$

Step 8. Calculate the second factor loadings. Use equation (5-8).

$$a_{i2} = \frac{r_{1,i8}}{\sqrt{r_{1,88}}} .$$ I-4

$$a_{12} = .11 \times 1.154 = .13$$
$$a_{22} = .05 \times 1.154 = .06$$

$$a_{32} = 0.$$
$$a_{42} = .40 \times 1.154 = .46.$$
$$a_{52} = .75 \times 1.154 = .87.$$

Enter the factor loadings in Column II of figure 9.

Check 3. $\displaystyle\sum_{i=1}^{5} a_{i2} \overset{?}{=} \frac{1}{\sqrt{r_{1,55}}} \sum_{i=1}^{5} r_{1,i5}.$ I-5

$$1.52 = 1.154 \times 1.32.$$
$$1.52 = 1.523.$$

Step 9. Repat the procedure commencing at step 4 until all factors have been extracted. For clarity, we shall continue the calculations using primes on the step numbers.

Step 4'. $\displaystyle\sum_{i=1}^{5} r_{2,ij} = \sum_{i=1}^{5} r_{1,ij} - a_{j2} \sum_{i=1}^{5} a_{i2}\ ;\ j = 1, \cdots, 5.$ I-6

$$\sum_i r_{2,i1} = 1.00 - .13 \times 1.52 = .80.$$
$$\sum_i r_{2,i2} = .71 - .06 \times 1.52 = .62.$$
$$\sum_i r_{2,i3} = 0.$$
$$\sum_i r_{2,i4} = 1.27 - .46 \times 1.52 = .57.$$
$$\sum_i r_{2,i5} = 1.32 - .87 \times 1.52 = 0.$$

Step 5'. Enter the column sums in column R2 of figure 10.

Step 6'. Variable 1 is the third pivot variable. Calculate the second residual correlations with variable 1. Enter them in the second column of figure 10.

$$r_{2,iq} = r_{iq} - \sum_{r=1}^{2} a_{ir} a_{qr}\ ;\ q = 1;\ i = 1, \cdots 5.$$ I-7

$$r_{2,11} = 1.00 - .54^2 - .13^2 = .69.$$
$$r_{2,21} = .50 - (.54 \times .71 + .06 \times .13) = .11$$
$$r_{2,31} = 0.$$
$$r_{2,41} = .33 - (.49 \times .54 + .46 \times .13) = .01$$
$$r_{2,51} = 0.$$

Check 1'. $\displaystyle\sum_{i=1}^{5} r_{2,i1}\ (\text{FIG. } 11) = .81\ ;\ \sum_{i=1}^{5} r_{2,i1}\ (\text{FIG } 10) = .80$

Check 2'. $\sum_i \sum_j r_{ij} - \sum_{T=1}^{2} \left(\sum_{i=1}^{5} a_{iT} \right)^2 \stackrel{?}{=} \sum_{ij} r_{2,ij}.$

$$14.80 - 10.50 - 2.31 = 1.99 \qquad \text{I-8}$$

$$1.99 = 1.99.$$

Step 7'. Calculate $\dfrac{1}{\sqrt{r_{2,11}}} = 1.204.$

Step 8'. Calculate the third factor loadings. Enter them in column III of figure 9.

$$a_{i3} = \frac{r_{2,iq}}{\sqrt{r_{2,qq}}}; \quad q = 1. \qquad \text{I-9}$$

$a_{13} = .69 \times 1.204 = .83.$

$a_{23} = .11 \times 1.204 = .13.$

$a_{33} = 0.$

$a_{43} = .01 \times 1.204 = .01.$

$a_{53} = 0.$

$\sum_{i=1}^{5} a_{i3} = .97$

Check 3'. $\sum_{i=1}^{5} a_{i3} = \dfrac{1}{\sqrt{r_{2,11}}} \sum_{i=1}^{5} r_{2,i1}$

$$.97 = 1.204 \times .81 = .975 \qquad \text{I-10}$$

Step 9'. Repeat the process starting at step 4. Double primes will indicate the second repetition.

Step 4''.

$$\sum_{i=1}^{5} r_{3,ij} = \sum_{i=1}^{5} r_{2,ij} - a_{j3} \sum_{i=1}^{5} a_{i3}. \qquad \text{I-11}$$

$\sum_i r_{3,i1} = .80 - .83 \times .97 = 00.$

$\sum_i r_{3,i2} = .62 - .13 \times .97 = .49.$

$\sum_i r_{3,i3} = 0.$

$\sum_i r_{3,i4} = .57 - .01 \times .97 = .56.$

$\sum_i r_{3,i5} = 0.$

Step 5''. Enter the column sums in column R_3 of figure 10.

Step 6''. Variable 4 is the fourth pivot variable. Calculate the residual correlations. Enter results in figure 11.

$$r_{3,is} = r_{is} - \sum_{T=1}^{3} a_{iT} \cdot a_{sT} \quad ; \quad i = 1, \cdots, 5; \; s = 4 \qquad \text{I-12}$$

$r_{3,14} = 0$

$r_{3,24} = .39 - (.71 \times .49 + .06 \times .46 + .13 \times .01) = .02$

$r_{3,34} = 0$

$r_{3,44} = 1.00 - (.49^2 + .46^2 + .01^2) = .55$

$r_{3,54} = 0$

$\sum_{i=1}^{5} r_{3,i4} = .57$

Check 1''. $\sum_{i=1}^{5} r_{3,i4}(\text{FIG. }10) = .57, \quad \sum_{i=1}^{5} r_{3,i4}(\text{FIG. }11) = .56$

Check 2''. $\sum_i \sum_j r_{ij} - \sum_{T=1}^{3} \left[\sum a_{iT}\right]^2 = \sum_i \sum_j r_{3,ij}.$ I-13

$14.80 - 10.50 - 2.31 - .94 \stackrel{?}{=} 1.05.$

$14.80 - 13.75 = 1.05$

Step 7''. Calculate $\dfrac{1}{\sqrt{r_{4,qq}}} = \dfrac{1}{\sqrt{.55}} = 1.334.$

Step 8''. Calculate the fourth factor loadings.

$$a_{i4} = \frac{r_{4,iq}}{\sqrt{r_{4,qq}}} \qquad \text{I-14}$$

$a_{14} = 0$

$a_{24} = .02 \times 1.335 = .03$

$a_{34} = 0$

$a_{44} = .55 \times 1.335 = .73$

$a_{54} = 0$

$\sum_{i=1}^{5} a_{i4} = .76$

Check 3''. $\sum_{i=1}^{5} a_{i4} \stackrel{?}{=} \dfrac{1}{\sqrt{r_{4,44}}} \sum_{i=1}^{5} r_{3,i4}$

$.76 = 1.335 \times .57$

$.76 = .7609$

Step 4'''.

$$\sum_{i=1}^{5} r_{4,ij} = \sum_{i=1}^{5} r_{3,ij} - a_{j4}\sum_{i=1}^{5} a_{i4}$$ I-15

$$\sum_{i=1}^{5} r_{4,i1} = 0$$

$$\sum_{i} r_{4,i2} = .49 - .03 \times .76 = .47$$

$$\sum_{i} r_{4,i3} = 0$$

$$\sum_{i} r_{4,i4} = .56 - .73 \times .76 = 0$$

$$\sum_{i} r_{4,i5} = 0$$

Step 5'''. Enter the column sums in column R_4 of figure 10.

Step 6'''. Variable 2 is the pivot variable. Calculate the residual correlations. Enter the results in figure 11.

$$r_{4,12} = 0$$ I-16

$$r_{4,22} = 1.00 - [(.71)^2 + (.06)^2 + (.13)^2 + (.03)^2] = 1.00 - .53 = .47$$

$$r_{4,32} = 0$$

$$r_{4,42} = 0$$

$$r_{4,52} = 0$$

$$\sum_{i=1}^{5} r_{4,i2} = .47$$

Check 1'''.

$$\sum r_{4,i2} \,|\, \text{FIG. 10} \,| = .47 = \sum_{i=1}^{5} r_{4,i2} \,|\, \text{FIG. 11} \,|.$$

Check 2'''. I-17

$$\sum_{i}\sum_{j} r_{ij} - \sum_{T=1}^{4}\left(\sum_{i=1}^{5} a_{iT}\right)^2 = \sum_{i}\sum_{j} r_{4,ij}$$

$$14.80 - 14.33 = .47$$

Step 8'''. Calculate the fifth factor loadings.

$$a_{i5} = \frac{r_{5,i2}}{\sqrt{r_{5,22}}}$$ I-18

$$a_{15} = 0$$

$$a_{25} = \sqrt{.47} = .69$$

$$a_{35} = 0$$

$$a_{45} = 0$$

$$a_{55} = 0.$$

Check 3'''. $$\sum a_{i5} = \frac{1}{\sqrt{r_{5,22}}}\sum_{i=1}^{5} r_{4,i4}$$

$$.69 = \frac{.47}{\sqrt{.47}} = .69$$

	1	2	3	4	5	
1	1.00	.50	.54	.33	.38	
2	.50	1.00	.71	.39	.41	
3	.54	.71	1.00	.49	.50	
4	.33	.39	.49	1.00	.65	
5	.38	.41	.50	.65	1.00	
$\sum_{i=1}^{5} r_{ij}$	2.75	3.01	3.24	2.86	2.94	$\sum_i \sum_j r_{ij} = 14.80$

Figure 8. The Complete Correlation Matrix

	I	II	III	IV	V
1	.54	.13	.83	0	0
2	.71	.06	.13	.03	.69
3	1.00	0	0	0	0
4	.49	.46	.01	.73	0
5	.50	.87	0	0	0
$\sum_{i=1}^{5} a_{ik}$	3.24	1.52	.97	.76	.69
$\left(\sum a_{ik}\right)^2$	10.50	2.31	.94	.58	.47

Figure 9. The Factor Loading Matrix

Residual Correlation Matrix

	R_1	R_2	R_3	R_4	R_5
1	1.00	.80	0	0	0
2	.71	.62	.49	.47	0
3	0	0	0	0	0
4	1.27	.57	.56	0	0
5	1.32	0	0	0	0
$\sum_i \sum_j r_{k,ij}$	4.30	1.99	1.05	.47	0

Figure 10. Column Sums of the Residual Correlation Matrices

Pivot Variables

	5	1	4	2
	R_1	R_2	R_3	R_4
1	.11	.69	0	0
2	.05	.11	.02	.47
3	0	0	0	0
4	.40	.01	.55	0
5	.75	0	0	0
$\sum_{i=1}^{5} r_{k,ij}$	1.31	.81	.57	.47

Figure 11. Residual Correlations with the Pivot Variables

APPENDIX II

THE COMPLETE CENTROID METHOD OF FACTORING

The following numerical example has been prepared in a step by step manner, paralleling, insofar as practical, the development in Chapter 6.[3] Where the example apparently does not exactly fit the theoretical development, the lack of fit is produced by simplifications in the computational technique. The example is started with a hypothetical correlation matrix and will demonstrate the extraction of the first and second centroid factors.

Variable	1	2	3	4	5	$\sum_{j=1}^{5} r_{ij};\ i \neq j$
1	(.54)	.50	.54	.33	.38	1.75
2	.50	(.71)	.71	.39	.41	2.01
3	.54	.71	(.71)	.49	.50	2.24
4	.33	.39	.49	(.65)	.65	1.86
5	.38	.41	.50	.65	(.65)	1.94
$\sum_{i=1}^{5} r_{ij};\ i \neq j$	1.75	2.01	2.24	1.86	1.94	
t_i	2.29	2.72	2.95	2.51	2.59	T = 13.06
						$\sqrt{T}$ = 3.613862
a_{i1}	.631	.753	.816	.695	.717	$\frac{1}{\sqrt{T}}$ = .276712

Figure 12. The Reduced Correlation Matrix

Step 1. Insert the communalities, determined by one of the methods described in Chapter 4. In this example, the communality in a particular column has been chosen to be equal to the largest correlation in that column.

Step 2. Add each column algebraically omitting the diagonal entries. If any of the column sums are negative, follow the reflection procedure described in step 11.

Check 1. Add each row and compare each row sum with the corresponding column total. They should be equal. In this example, it may seem trivial to perform this check; however, it is strongly recommended as a means of reducing errors.

Step 3. Add each communality to the appropriate column total. This sum is t_j defined in equation 6-8.

Step 4. Total the t_j to form T.

Step 5. Compute the square root of T and its reciprocal.

Step 6. Compute the a_{j1} for all j, employing equation 6-8.

Sample calculation

$$a_{11} = \frac{2.29}{\sqrt{13.06}} = .634. \qquad \text{II-1}$$

Check 2. Does the sum of the a_j for all j equal the square root of T?

$$\sum_j a_{j1} = 3.615 \; ; \sqrt{T} = 3.613862. \qquad \text{II-2}$$

Step 7. Place the factor loadings vertically and in order in the first column of the factor loading matrix. See figure 13. If any of the variables have been reflected, the signs of the factor loadings on those variables are negative. The signs of those variables which have not been reflected are positive.

Step 8. Prepare a matrix form and enter the information which is typewritten in figure 14. Although there may seem to be a considerable amount of work in preparing the matrix, it is recommended as a means of preventing errors.

Step 9. Compute all of the elements, $r_{1,ij}$ by applying equation 6-9.

$$r_{1,11} = .54 - .634 \times .634 = .138. \qquad \text{II-3}$$

Check 3. Add each column and row. Are the totals equal to zero within rounding errors? If so, the work is probably correct to this point. See figure 14.

	FACTOR	
	I	II
1	.634	.219
2	.753	.329
3	.816	.241
4	.695	-.399
5	.717	-.379

Figure 13. The Centroid Factor Loading Matrix

		.634	.753	.816	.695 *	.717 *	$\sum$ Row
		1	2	3	4	5	
.634	1	.54	.50	.54	.33	.38	
		.111 $\boxed{.138}$	.023	.023	±.111	±.075	-.002
.753	2	.50	.71	.71	.39	.41	
		.023	.133 $\boxed{.143}$	.096	±.133	±.130	-.001
.816	3	.54	.71	.71	.49	.50	
		.023	.096	.044	±.077	±.085	.001
* .695	4	.33	.39	.49	.65	.65	
		±.111	±.133	±.077	.152 $\boxed{.167}$	.152	-.002
.717	5	.38	.41	.50	.65	.65	
		±.075	±.130	±.085	.152	.152 $\boxed{.136}$	-.002
$\sum$ Column		-.002	-.001	.001	-.002	-.002	
$\sum_i r_{1ij}$ $i \neq j$		-.140	-.144	-.043	-.169	-.138	$\sum\sum = -.634$
Col. 4		.085	.122	.111	.169	-.442	
Col. 5		.232	.382	.281	.473	.442	
t_{i2}		.343	.515	.377	.625	.594	$T_2 = 2.454$
a_{j2}		.219	.329	.241	.399	.379	$\sqrt{T_2} = 1.566524$ $\frac{1}{\sqrt{T_2}} = .638356$

Figure 14. The First Residual Correlation Matrix

Step 10. Estimate the communalities in the first residual correlation matrix as in step 1. Replace the entries in figure 14 with the new communalities. They are indicated in the figure by underlines.

Step 11. Repeat step 2. Note that all of the column sums are negative. Therefore we must reflect some or all of the variables in order to proceed with the centroid method. Since it is desired to move the centroid as far from the origin as possible, we want to make the total algebraic sum of all of the values in the matrix as large a positive value as possible. This may be accomplished as follows:

a. Select the column with the largest negative total in the residual correlation matrix. In our example, column 4 in figure 14 has the largest negative sum. Copy the column total with sign reversed in a row immediately below the row of column sums. Label this row with the number of the column being reflected.

b. Place an asterisk at the top of the column being reflected and in front of the corresponding row to indicate that the variable has been reflected.

c. In each of the other columns, double the value of the residual in the row marked with an asterisk, change its sign and add that amount to the previous column total. Enter each new total in the row marked "Column 4" in the appropriate column.

Example: the value in the second column of the row marked "Column 4" is computed as follows:

$$-.144 + 2(.133) = .122 \qquad \text{II-4}$$

Check 3. When all of the values in the row marked "Column 4" have been computed, they should be totaled. That total should equal the total of the original column totals plus four times the column sum of the variable

which has been reflected.

$$-.634 + 4(.169) = .042 \qquad \text{II-5}$$

Step 11. d. If any of the new column totals (those in the row marked "Column 4") are negative, select the variable with the largest negative sum and reflect that variable. In our example, column 5 should be reflected. In columns which have been previously reflected once, the signs are not reversed before adding the doubled values. It may happen that it is necessary to reflect a variable more than once in the same residual correlation matrix. For the first, third, or any odd number of reflections, the sign of the doubled value must be changed before adding to the previous column total. For even numbered reflections, the sign of the doubled value is not changed.

e. Repeat the reflection process until all column totals are either zero or positive. Check 3 should be applied on each new row of column totals to guard against error.

Step 12. Change the signs of all of the reflected row values which are not in reflected columns and reverse the signs of all reflected column values not in reflected rows. (See figure 13.) The entries which are marked with both a plus and minus sign were originally of the sign indicated in the lower position. That is, if $\pm$ then the sign was originally negative and if $\mp$, then the sign was positive.

Step 13. Compute the factor loadings for the next factor:

a. Add to the last row of adjusted column sums the communality for each column. These sums are located in the row, t_{i2}.

b. Add the t_{i2} to form the total sum T_2.

c. The factor loading for variable i on the second factor is II-6

Step 14. Place the loadings in column II of figure 13.

Step 15. Determine the signs of the loadings.

a. The sign of a loading on a variable that has been reflected an odd number of times is given the sign opposite to the sign of the loading on the previous factor for the same variable.

b. The sign of a factor loading for a variable which has not been reflected or which has been reflected an even number of times is given the same sign as the previous factor loading.

The process will be repeated until one of two results is produced: the number of factors desired has been extracted; or, the residual correlations become so small that they are no longer significant when compared to rounding errors. The acid test regarding the number of factors is to attempt to reconstruct the correlation matrix by multiplying the factor loading matrix by its inverse. If the correlation matrix is reproduced to an accuracy compatible with the rounding errors, then the number of factors and the factor loadings certainly will be adequate.

BIBLIOGRAPHY

1. Cattall, R. B., Factor Analysis, Harper Brothers, 1952.

2. Creasy, M. A., Analysis of Variance as an Alternative to Factor Analysis, J. Roy. Stat. Soc., 1952, 19, pp. 318-325.

3. Fruchter, B., Introduction to Factor Analysis, Van Nostrand, 1954.

4. Holzinger, K. J. and Harman, H. H., Factor Analysis, University of Chicago Press, 1941.

5. Kendall, M. G. and Babbington Smith, B., Factor Analysis, J. Roy. Stat. Soc. Series B, 12, pp. 60-94.

6. Rao, C. R., Estimation and Tests of Significance in Factor Analysis, Psychometrika, 1955, 20, pp. 93-109.

7. Thompson, G. H., The Factorial Analysis of Human Ability, University of London Press, 5th Edition, 1951.

8. Thurstone, L. L., Multiple Factor Analysis, University of Chicago Press, 1947.

9. ________, Vectors of the Mind, University of Chicago Press, 1935.

10. Vernon, P. E., The Structure of Human Abilities, Wiley, 1950.

11. Statistical Methodology in Educational Research, Review of Educational Research, XXIV, No. 5, 1954.

Notes: 1. Two publications contain a great deal of information relative to factor analysis. They are Psychometrika and The British Journal of Psychology, Statistical Section.

2. The reports published under AF Contract 33(038)-25726 by the University of Illinois contain rather complete bibliographies of the mathematical and statistical references in the theory of factor analysis as of 1953.